The Pareto Regret Frontier

Wouter M. Koolen
The online learning philosophy

Want to solve a learning problem?
The online learning philosophy

Want to solve a learning problem?
- Get hold of a bunch of
  - trivial approaches 😊
The online learning philosophy

Want to solve a learning problem?

▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊

▶ Each “expert” implements a solution
▶ Use aggregation algorithm to combine solutions in production
▶ Profit!
The online learning philosophy

Want to solve a learning problem?

▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊
  ▶ students ($)
The online learning philosophy

Want to solve a learning problem?
▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊
  ▶ students ($)
  ▶ employees ($$)

Profit!
The online learning philosophy

Want to solve a learning problem?

▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊
  ▶ students ($) 
  ▶ employees ($$)
  ▶ companies ($$$)
The online learning philosophy

Want to solve a learning problem?

- Get hold of a bunch of
  - trivial approaches 😊
  - models from literature 😐
  - students ($)
  - employees ($$)
  - companies ($$$)
  - ...

Use aggregation algorithm to combine solutions in production

Profit!
The online learning philosophy

Want to solve a learning problem?

▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊
  ▶ students ($)  
  ▶ employees ($$)
  ▶ companies ($$$)
  ▶ . . .

▶ Each “expert” implements a solution
The online learning philosophy

Want to solve a learning problem?

▸ Get hold of a bunch of
  ▸ trivial approaches 😊
  ▸ models from literature 😊
  ▸ students ($)
  ▸ employees ($$)
  ▸ companies ($$$)
  ▸ . . .

▸ Each “expert” implements a solution

▸ Use aggregation algorithm to combine solutions in production
The online learning philosophy

Want to solve a learning problem?

▶ Get hold of a bunch of
  ▶ trivial approaches 😊
  ▶ models from literature 😊
  ▶ students ($) 
  ▶ employees ($$)
  ▶ companies ($$$)
  ▶ ...

▶ Each “expert” implements a solution
▶ Use aggregation algorithm to combine solutions in production

▶ Profit!
That aggregation algorithm

\( T \) rounds, \( K \) experts

\[ L_T - \min_k L_T^k \leq \sqrt{\frac{T}{2 \ln K}} \]

regret
That aggregation algorithm

\( T \) rounds, \( K \) experts

\[
L_T - \min_k L_T^k \leq \sqrt{\frac{T}{2} \ln K}
\]

regret

What if

- Lots of experts? shotgun-style “throw in all you got”
- Special experts? company’s current strategy
Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is $T$-realisable if there is a strategy ensuring

$$L_T - L^k_T \leq r_k$$

for each expert $k$.
Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is $T$-realisable if there is a strategy ensuring

$$L_T - L_T^k \leq r_k$$

for each expert $k$

Suggestive: for every expert prior $\mathbb{P}$, the following is $T$-realisable:

$$\left\langle \sqrt{\frac{T}{2} \left( - \ln \mathbb{P}(k) \right)} \right\rangle_{k=1}^K$$

but this is false

So what can be realised?
Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is $T$-realisable if there is a strategy ensuring

$$L_T - L^k_T \leq r_k \text{ for each expert } k$$

Suggestive: for every expert prior $\mathbb{P}$, the following is $T$-realisable:

$$\left\langle \sqrt{\frac{T}{2}} \left( - \ln \mathbb{P}(k) \right) \right\rangle_{k=1}^K$$

but this is false
Regret as a multi-objective criterion

A vector $\langle r_1 \ldots r_K \rangle$ is $T$-realisable if there is a strategy ensuring

$$L_T - L_T^k \leq r_k$$ for each expert $k$.

Suggestive: for every expert prior $\mathbb{P}$, the following is $T$-realisable:

$$\left\langle \sqrt{\frac{T}{2}} \left( - \ln \mathbb{P}(k) \right) \right\rangle_{k=1}^K$$

but this is false.

So what can be realised?
Results in a nutshell

- Absolute loss (or $K = 2$ experts)
  - Exact results
    - Characterisation of $T$-realisable frontier (combinatorial)
    - Strategy for each trade-off

- Asymptotic (large $T$):
  - Smooth limit frontier
  - Smooth limit strategy
  - For any $p \in [0, 1]$ we can realise:
    $$\sqrt{T} \left( -\ln(p) \right), \sqrt{T} \left( -\ln(1-p) \right)$$
    but we can do better(!)

- For $K > 2$ experts
  - For every expert prior $P(k)$, we can realise
    $$\sqrt{2.6T} \left( -\ln P(k) \right)$$
    using a recursive combination of 2-expert algorithms.
Results in a nutshell

- Absolute loss (or $K = 2$ experts)
  - Exact results
    - Characterisation of $T$-realisable frontier (combinatorial)
    - Strategy for each trade-off
  - Asymptotic (large $T$):
    - Smooth limit frontier
    - Smooth limit strategy
    - For any $p \in [0, 1]$ we can realise:
      \[
      \left\langle \sqrt{T \left( -\ln(p) \right)}, \sqrt{T \left( -\ln(1 - p) \right)} \right\rangle
      \]
      but we can do better(!)
Results in a nutshell

- Absolute loss (or \( K = 2 \) experts)
  - Exact results
    - Characterisation of \( T \)-realisable frontier (combinatorial)
    - Strategy for each trade-off
  - Asymptotic (large \( T \)):
    - Smooth limit frontier
    - Smooth limit strategy
    - For any \( p \in [0, 1] \) we can realise:
      \[
      \left( \sqrt{T (-\ln(p))}, \sqrt{T (-\ln(1-p))} \right)
      \]
      but we can do better(!)

- For \( K > 2 \) experts
  - For every expert prior \( \mathbb{P}(k) \), we can realise
    \[
    \left( \sqrt{2.6 T (-\ln \mathbb{P}(k))} \right)_{k=1}^{K}
    \]
    using a recursive combination of 2-expert algorithms.
Absolute loss game

Each round $t \in \{1, \ldots, T\}$ the learner assigns a probability $p_t \in [0, 1]$ to the next outcome being a 1, after which the actual outcome $x_t \in \{0, 1\}$ is revealed, and the learner suffers

$$\text{absolute loss} \quad |p_t - x_t|.$$  

The regret w.r.t. the strategy that always predicts $k \in \{0, 1\}$ is

$$R^k_T := \sum_{t=1}^{T} (|p_t - x_t| - |k - x_t|)$$

A candidate trade-off $\langle r_0, r_1 \rangle \in \mathbb{R}^2$ is called $T$-realisable for the $T$-round absolute loss game if there is a strategy that keeps the regret w.r.t. each $k \in \{0, 1\}$ below $r_k$, i.e. if

$$\exists p_1 \forall x_1 \cdots \exists p_T \forall x_T : R^0_T \leq r_0 \text{ and } R^1_T \leq r_1$$

We denote the set of all $T$-realisable trade-offs by $\mathcal{G}_T$. 
The set $\mathcal{G}_T$, i.e. the $T$-realisable tradeoffs $\langle r_0, r_1 \rangle$
Theorem
The Pareto frontier of $G_T$ is piece-wise linear with $T + 1$ vertices:

\[ \langle f_T(i), f_T(T - i) \rangle \quad 0 \leq i \leq T \quad \text{where} \quad f_T(i) := \sum_{j=0}^{i} j2^{j-T} \left( \frac{T - j - 1}{T - i - 1} \right). \]

The optimal strategy at vertex $i$ assigns to $x = 1$ probability

\[ p_T(0) := 0, \quad p_T(T) := 1, \quad p_T(i) := \frac{f_{T-1}(i) - f_{T-1}(i - 1)}{2} \quad 0 < i < T, \]

and it interpolates linearly in between consecutive vertices.
Asymptotic analysis

Idea: normalise and then make $T$ large
Asymptotic analysis

Idea: normalise and then make $T$ large

$$G := \lim_{T \to \infty} \frac{G_T}{\sqrt{T}}.$$
Asymptotic plot (moderate)

Limit frontier: $T = 10000$ and $\sqrt{-\ln \mathbb{P}(k)}$,
Asymptotic plot (tail)

normalised regret w.r.t. expert 2

limit frontier

$T = 10000$

$\sqrt{-\ln \mathbb{P}(k)}$

normalised regret w.r.t. expert 1
Asymptotic Pareto Frontier and Optimal Strategy

Theorem

The Pareto frontier of $\mathbb{G}$ is the smooth curve

$$\langle f(u), f(-u) \rangle \quad u \in \mathbb{R}, \quad \text{where} \quad f(u) := \int_{-\infty}^{u} \Phi(x) \, dx,$$

and $\Phi(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{x^2}{2}} \, dx$ is the standard normal CDF. The optimal strategy converges to

$$p(u) = \Phi(u).$$
Use of asymptotics

- Smooth formula easier to handle
- Allows us to appreciate that sqrt-min-log-prior tradeoffs are realisable with constant 1 (not $1/\sqrt{2}$) …
- …but have suboptimal lower-order terms.
- Suggests smoothened algorithms
Combine algorithm for $K = 2$ into unbalanced binary tree.

Outermost algorithm combines expert with least prior vs rest

Gives us

$$\sqrt{2.6 T \left( - \ln \mathbb{P}(k) \right)}$$
Combine algorithm for $K = 2$ into unbalanced binary tree.

Outermost algorithm combines expert with least prior vs rest

Gives us

$$\sqrt{2.6 \mathcal{T} \left( -\ln \mathbb{P}(k) \right)}$$

But we would like to have the exact Pareto frontier.
Conclusion

- We need unfair regret bounds
- Reinterpret regret as a multi-criterion objective
- Exact Pareto frontier for $K = 2$ experts
- with optimal algorithm
- And useful formula for asymptotic Pareto frontier
- with asymptotic algorithm
- Trick for $K > 2$ experts